

**REMARKS**

Enclosed herewith is a supplemental Information Disclosure Statement listing the reference C. Manning, et al., Foundations of Statistical Natural Language Processing, MIT Press, Cambridge, MA, 2000, pp. 539-544 and a copy of the reference. This reference was cited in a previously filed Information Disclosure Statement and a copy of the reference was supplied to the Examiner, however, that previously supplied copy was missing page 544. The copy enclosed herewith is a complete copy of the reference including pp. 539-544. Applicant respectfully requests that this reference be entered.

On page 2 of the Office Action, claims 1-6, 8-14 and 16 were rejected under 35 U.S.C. § 103 as being unpatentable over U.S. Patent No. 6,026,388 to Liddy et al. (hereinafter "Liddy") in further view of U.S. Patent Application No. 2002/0002479 filed by Almog et al. (hereinafter "Almog").

**Description of the Present Invention**

The present invention provides an approach for determining whether records stored in the same database that contain different types of data, such as structured and unstructured data, are similar. In accordance with the invention, multiple records in the database are accessed and corresponding fields in the records evaluated to determine a data type associated with data contained in the fields. An appropriate matching process is then selected depending on the data type. The matching process selected is one of a Boolean matching process, an ordinal matching process or a vector-based matching process. The selected matching process is then applied to determine a match score for data contained in the corresponding fields.

**Description of Cited References**

Liddy describes a technique for representing queries and documents in a retrieval system using natural language processing (NLP) to enable representing, indexing and retrieving text at multiple levels. See Col. 2, lines 35-41. According to Liddy, documents are processed to generate an alternative representation for the documents. This alternative presentation is used to

match queries with documents. See Col. 7, lines 29-41. The processing involves a series of stages which are applied to each document to generate the alternative representation for the document. See Col. 7, lines 29-31, Col. 8, line 49 to Col. 9, line 15 and Fig. 3. The alternative representation is stored in a document database. See Col. 7, lines 14-21 and Col. 8, lines 18-21, see also, Fig. 1.

Likewise, the system processes queries including generating an alternative representation for the queries. See Col. 2, lines 46-48. This processing involves a set of modules which process the query stage-by-stage to generate the alternative representation. See Col. 16, line 34 to Col. 17, line 18 and Fig. 5. The system matches the alternative representation of a query to alternative representation of documents contained in the document database to obtain a measure of relevance between the query and documents in the database. See Col. 2, lines 61-67, see also, Col. 8, lines 29-48 and Col. 32, line 36 to Col. 33, line 6.

Almog describes a job placement system that involves two databases, a job opening database and a worker database. See Abstract. The job opening database contains records that hold descriptions of job openings and the worker database contains records that hold profiles of workers. A worker is matched to one or more job openings by comparing a record holding the worker's profile stored in the worker database with records holding job openings stored in the job database to determine if a listing of job openings match the profile. See paragraphs 0018, 0040, and Figs. 2 and 3.

#### Differences between the Present Invention and the Cited Art

Applicant respectfully urges that Liddy or Almog taken either singly or in combination do not render Applicant's claimed invention unpatentable under 35 U.S.C. § 103.

Representative claim 1 recites in relevant part:

1. A method for determining whether records are similar in a database containing both structured and unstructured, free-text data, the method comprising the steps of:

*accessing two of the records from the database* for evaluation; and  
*evaluating a match between the two records* as a weighted match between each of a plurality of available fields, *such that a matching process is selected* as appropriate from among a group of matching processes including strict Boolean, ordinal, and vector-based matching processes, ...

Applicant submits that neither Liddy nor Almog teach or suggest Applicant's claimed *accessing two records in a database and evaluating a match between the two records such that a matching process is selected*. Both Liddy and Almog are silent with regards to these two limitations. At best Liddy describes generating alternative representations for documents and queries before matching documents with queries. These alternative representations are generated by submitting the queries and documents to various stages of processing that form the alternative representations. These stages of processing are cited by the Examiner as matching processes. However, Applicant respectfully disagrees. Applicant submits that these stages are used to generate alternative representations for queries and documents and are not used by the matching process. Rather, matching occurs after alternative representations have been generated for queries and documents. These alternative representations are then used to match queries with documents to determine a degree of relevance between the documents and the queries.

Applicant's claimed invention, on the other hand, does not require generating an alternative representation of data before matching occurs. Rather, Applicant's claimed invention obviates the need to generate an alternative representation because a matching process is selected based on the type of data being matched in records at the time the data is being matched.

Regarding Almog, Almog teaches comparing records in one database with records in another database. Applicant's claimed invention, on the other hand, claims comparing records in the same database and does not require two databases as is taught by Almog.

For reasons set forth above, Applicant respectfully urges that neither Liddy nor Almog taken either singly or in combination render claims 1-6, 8-14 and 16 in Applicant's application unpatentable under 35 U.S.C. §103 and therefore believes claims 1-6, 8-14 and 16 are in condition for allowance.

At paragraph 7 of the Office Action, claims 8 and 16 were rejected under 35 U.S.C. § 103 as being unpatentable over Liddy, Almog and in further view of U.S. Patent 5,778,388 to Kawamura et al. (hereinafter "Kawamura").

Kawamura describes a technique for retaining throughput of a transaction process during a syncpoint acquisition process, reducing the time elapsed in a system restart process when restarting a system operation and simultaneously accessing a database's updated page existing at

the initiation of a syncpoint acquisition and a page that is not updated in a buffer pool even during syncpoint acquisition. See Col. 3, lines 44-55. According to the technique, a mark indicating that a syncpoint acquisition is in progress is stored in a table to obtain a syncpoint in a database to enable restarting and restoring a transaction in the event of a system failure. Pages (updated pages) to be written to the database are marked. During the syncpoint acquisition process, the marked pages are written to the database. As each page is written to the database, it's mark is released (thus the page may be accessed). The syncpoint acquisition process is considered finished when all of the marked pages are written to the database. Releasing the mark associated with marked pages after the page has been written enables a transaction process to be continued during the syncpoint acquisition process. See Col. 3, line 56 to Col. 4, line 7.

Applicant respectfully submits that Liddy, Almog and Kawamura taken either singly or in combination do not teach Applicant's claimed accessing two records from a database for evaluation and evaluating a match between the two records as a weighted match between each of a plurality of available fields, such that a matching process is selected as appropriate from among a group of matching processes including strict Boolean, ordinal, and vector-based matching processes. Therefore, Applicant respectfully believes that claims 8 and 16 are in condition for allowance.

All independent claims are believed to be in condition for allowance.

All dependent claims are believed to be dependent on believed to be allowable independent claims and therefore in condition for allowance.

Quick favorable action is respectfully requested.

**CONCLUSION**

In view of the above amendments and remarks, it is believed that all claims are in condition for allowance, and it is respectfully requested that the application be passed to issue. If the Examiner feels that a telephone conference would expedite prosecution of this case, the Examiner is invited to call the undersigned.

Respectfully submitted,

HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

By 

David J. Thibodeau, Jr.

Reg. No. 31,671

Telephone: (978) 341-0036

Facsimile: (978) 341-0136

Concord, MA 01742-9133

Dated: 11/20/05